SEMICONDUCTOR DEVICE AND METHOD OF MANUFACTURING THE SAME

Japanese Patent Application No. 2002-309839, filed on October 24, 2002, is hereby incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

The present invention relates to a semiconductor device having a memory region and a method of manufacturing the semiconductor device, and in particular, to a semiconductor device in which a non-volatile memory device formed within the memory region includes two charge accumulation regions for each word gate, and a method of manufacturing the semiconductor device.

One type of non-volatile memory device is called a metal-oxide-nitride-oxide semiconductor (MONOS) type or a silicon-oxide-nitride-oxide-silicon (SONOS) type, wherein a gate insulating layer between a channel region and a control gate is formed of a multi-layer stack of silicon oxide and silicon nitride layers, and charge is trapped in the silicon nitride layer.

A device shown in Fig. 17 is known as an example of this MONOS type of non-volatile memory device (non-patent document by Y. Hayashi, et al, in 2000 Symposium on VLSI Technology Digest of Technical Papers, pp. 122-123).

In this MONOS memory cell 100, a word gate 14 is formed on a semiconductor substrate 10 with a gate insulating layer 12 therebetween. A control gate 20 and a control gate 30 are disposed on either side of the word gate 14, in the shape of side walls. There is a first insulating layer 22 between a base portion of the control gate 20 and the semiconductor substrate 10, and a side insulating layer 26 between a side surface of the control gate 20 and the word gate 14. In a similar manner, the first insulating layer 22 is also between a base portion of the control gate 30 and the semiconductor substrate 10, and the side insulating layer 26 is also between a side

surface of the control gate 30 and the word gate 14. Impurity layers 16 and 18, which are to form a source region and drain region, are formed in the semiconductor substrate 10 between the opposing control gates 20 and 30 of neighboring memory cells.

In this manner, each memory cell 100 has two MONOS memory elements on the side surfaces of the word gate 14. These two MONOS memory elements can be controlled independently. Thus one memory cell 100 can store two bits of information.

BRIEF SUMMARY OF THE INVENTION

The present invention may provide a semiconductor device which includes MONOS type non-volatile memory devices, each having two charge accumulation regions, and particularly a semiconductor device having resistance to deterioration during the writing/erasing of data, and a method of manufacturing thereof.

According to a first aspect of the present invention, there is provided a semiconductor device having a memory region in which a memory cell array is formed of non-volatile memory devices arranged in a matrix of a plurality of rows and columns,

wherein each of the non-volatile memory devices has:

a word gate formed above a semiconductor layer with a gate insulating layer interposed;

an impurity layer formed in the semiconductor layer to form a source region or a drain region; and

control gates in the form of side walls formed along both side surfaces of the word gate;

wherein each of the control gates consists of a first control gate and a second control gate adjacent to each other;

wherein a first insulating layer is disposed between the first control gate and the semiconductor layer, and a side insulating layer is disposed between the first control gate and the word gate;

wherein a second insulating layer is disposed between the second control gate and the semiconductor layer;

wherein the thickness of the second insulating layer is less than the thickness of the first insulating layer; and

wherein an uppermost layer of the second insulating layer is a charge transfer protection film.

According to a second aspect of the present invention, there is provided a method of manufacturing a semiconductor device having a memory region in which a memory cell array is formed of non-volatile memory devices arranged in a matrix of a plurality of rows and columns, the method comprising:

(a) forming a gate insulating layer above a semiconductor layer;

(b) forming a first conductive layer above the gate insulating layer;

(c) forming a stopper layer above the first conductive layer;

(d) patterning the stopper layer and the first conductive layer to form a stack of layers formed of that stopper layer and that first conductive layer;

(e) forming a first insulating layer by stacking a first silicon oxide film, a silicon nitride film, and a second silicon oxide film over the entire surface of the memory region;

(f) forming a second conductive layer above the first insulating layer, and then anisotropically etching the second conductive layer into side-wall-shaped first control gates on both side surfaces of the first conductive layer and on the semiconductor layer with the first insulating layer interposed;

(g) using the first control gate as a mask to remove a surface portion of the second silicon oxide film of the first insulating layer, and defining part of the remaining first insulating layer as a second insulating layer;

(h) forming a third conductive layer over the entire surface of the memory region, and then anisotropically etching the third conductive layer into a second control

gate on a side surface of each of the first control gates and on the semiconductor layer with the second insulating layer interposed;

(i) forming an impurity layer in the semiconductor layer to form a source region or a drain region;

(j) forming a third insulating layer over the entire surface of the memory region and then removing part of the third insulating layer to expose part of the stopper layer; and

(k) removing the stopper layer, forming a fourth conductive layer over the entire surface of the semiconductor layer, and then patterning the fourth conductive layer to form a word line.

According to a third aspect of the present invention, there is provided a method of manufacturing a semiconductor device having a memory region in which a memory cell array is formed of non-volatile memory devices arranged in a matrix of a plurality of rows and columns, the method comprising:

(a) forming a gate insulating layer above a semiconductor layer;

(b) forming a first conductive layer above the gate insulating layer;

(c) forming a stopper layer above the first conductive layer;

(d) patterning the stopper layer and the first conductive layer to form a stack of layers formed of that stopper layer and that first conductive layer;

(e) forming a first insulating layer by stacking a first silicon oxide film, a silicon nitride film, and a second silicon oxide film over the entire surface of the memory region;

(f) forming a second conductive layer above the first insulating layer, and then anisotropically etching the second conductive layer into side-wall-shaped first control gates on both side surfaces of the first conductive layer and on the semiconductor layer with the first insulating layer interposed;

(g) using the first control gate as a mask to remove part of the second silicon

oxide film of the first insulating layer and expose part of the silicon nitride film of the first insulating layer, forming a charge transfer protection film on the exposed portion of the silicon nitride film of the first insulating layer, and then defining part of the remaining first insulating layer and the charge transfer protection film as a second insulating layer;

(h) forming a third conductive layer over the entire surface of the memory region, and then anisotropically etching the third conductive layer into a second control gate on a side surface of each of the first control gates and on the semiconductor layer with the second insulating layer interposed;

(i) forming an impurity layer in the semiconductor layer to form a source region or a drain region;

(j) forming a third insulating layer over the entire surface of the memory region and then removing part of the third insulating layer to expose part of the stopper layer; and

(k) removing the stopper layer, forming a fourth conductive layer over the entire surface of the semiconductor layer, and then patterning the fourth conductive layer to form a word line.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

FIG. 1 is a plan view schematically showing a layout of a memory region of the semiconductor device.

FIG. 2 is a schematic cross-sectional view taken along the line A-A shown in FIG. 1.

FIG. 3 is a cross-sectional view schematically showing a portion B shown in FIG. 2.

FIG. 4 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 5 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 6 is a plan view showing a step of a method of manufacturing the semiconductor device shown in FIG. 5.

FIG. 7 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 8 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 9 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 10 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 11 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 12 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 13 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 14 is a cross-sectional view showing a step of a method of manufacturing the semiconductor device shown in FIGS. 1 to 3.

FIG. 15 is a cross-sectional view schematically showing a semiconductor device according to a second embodiment.

FIG. 16 is a cross-sectional view schematically showing a semiconductor device according to the second embodiment.

FIG. 17 is a cross-sectional view showing a conventional MONOS memory cell.

FIG. 18 is a diagram for illustrating an erase operation of a semiconductor

device according to the present invention.

DETAILED DESCRIPTION OF THE EMBODIMENTS

Embodiments of the present invention are described below.

According to one embodiment of the present invention, there is provided a semiconductor device having a memory region in which a memory cell array is formed of non-volatile memory devices arranged in a matrix of a plurality of rows and columns,

wherein each of the non-volatile memory devices has:

a word gate formed above a semiconductor layer with a gate insulating layer interposed;

an impurity layer formed in the semiconductor layer to form a source region or a drain region; and

control gates in the form of side walls formed along both side surfaces of the word gate;

wherein each of the control gates consists of a first control gate and a second control gate adjacent to each other;

wherein a first insulating layer is disposed between the first control gate and the semiconductor layer, and a side insulating layer is disposed between the first control gate and the word gate;

wherein a second insulating layer is disposed between the second control gate and the semiconductor layer;

wherein the thickness of the second insulating layer is less than the thickness of the first insulating layer; and

wherein an uppermost layer of the second insulating layer is a charge transfer protection film.

According to this semiconductor device, each control gate consists of the first control gate and the second control gate, and the first control gate and the second

control gate are respectively formed on the insulating layers having different thicknesses. Therefore, a semiconductor device in which the potential of the surface of the substrate under the control gates changes at two locations can be provided.

An uppermost layer of the second insulating layer is the charge transfer protection film. Therefore, the silicon nitride film which accumulates a charge can be prevented from coming in contact with the control gate. This prevents diffusion of the charge accumulated in the silicon nitride film into the second control gate, whereby the charge retention characteristics can be improved.

The semiconductor device in accordance with this embodiment could have following features.

(A) In this semiconductor device, the first insulating layer may be a stack of a first silicon oxide film, a silicon nitride film, and a second silicon oxide film.

(B) In this semiconductor device, the second insulating layer may be a stack of a silicon oxide film, a silicon nitride film and the charge transfer protection film, the thickness of the charge transfer protection film being less than the thickness of the second silicon oxide film of the first insulating layer.

(C) In this semiconductor device, the charge transfer protection film may be further provided on a surface of the first control gate.

(D) In this semiconductor device, the charge transfer protection film may be one of a silicon oxide film and a silicon oxide nitride film.

According to another embodiment of the present invention, there is provided a method of manufacturing a semiconductor device having a memory region in which a memory cell array is formed of non-volatile memory devices arranged in a matrix of a plurality of rows and columns, the method comprising:

(a) forming a gate insulating layer above a semiconductor layer;

(b) forming a first conductive layer above the gate insulating layer;

(c) forming a stopper layer above the first conductive layer;

(d) patterning the stopper layer and the first conductive layer to form a stack of layers formed of that stopper layer and that first conductive layer;

(e) forming a first insulating layer by stacking a first silicon oxide film, a silicon nitride film, and a second silicon oxide film over the entire surface of the memory region;

(f) forming a second conductive layer above the first insulating layer, and then anisotropically etching the second conductive layer into side-wall-shaped first control gates on both side surfaces of the first conductive layer and on the semiconductor layer with the first insulating layer interposed;

(g) using the first control gate as a mask to remove a surface portion of the second silicon oxide film of the first insulating layer, and defining part of the remaining first insulating layer as a second insulating layer;

(h) forming a third conductive layer over the entire surface of the memory region, and then anisotropically etching the third conductive layer into a second control gate on a side surface of each of the first control gates and on the semiconductor layer with the second insulating layer interposed;

(i) forming an impurity layer in the semiconductor layer to form a source region or a drain region;

(j) forming a third insulating layer over the entire surface of the memory region and then removing part of the third insulating layer to expose part of the stopper layer; and

(k) removing the stopper layer, forming a fourth conductive layer over the entire surface of the semiconductor layer, and then patterning the fourth conductive layer to form a word line.

According to this method of manufacturing a semiconductor device, the control gate is formed in two steps. In more detail, the first control gate is formed on the first insulating layer. A surface portion of the second silicon oxide film is then removed,

and part of the remaining first insulating layer is defined as the second insulating layer. The second control gate is formed on this second insulating layer. Therefore, the control gates can be respectively formed on the insulating layers having different thickness. As a result, a semiconductor device in which field intensity between the control gates and the surface of the substrate is nonuniform can be manufactured.

Moreover, the second silicon oxide film which partially remains on the silicon nitride film by removing only a surface portion of the second silicon oxide film in the step (g) functions as a charge transfer protection film.

According to a further embodiment of the present invention, there is provided a method of manufacturing a semiconductor device having a memory region in which a memory cell array is formed of non-volatile memory devices arranged in a matrix of a plurality of rows and columns, the method comprising:

(a) forming a gate insulating layer above a semiconductor layer;

(b) forming a first conductive layer above the gate insulating layer;

(c) forming a stopper layer above the first conductive layer;

(d) patterning the stopper layer and the first conductive layer to form a stack of layers formed of that stopper layer and that first conductive layer;

(e) forming a first insulating layer by stacking a first silicon oxide film, a silicon nitride film, and a second silicon oxide film over the entire surface of the memory region;

(f) forming a second conductive layer above the first insulating layer, and then anisotropically etching the second conductive layer into side-wall-shaped first control gates on both side surfaces of the first conductive layer and on the semiconductor layer with the first insulating layer interposed;

(g) using the first control gate as a mask to remove part of the second silicon oxide film of the first insulating layer and expose part of the silicon nitride film of the first insulating layer, forming a charge transfer protection film on the exposed portion of

the silicon nitride film of the first insulating layer, and then defining part of the remaining first insulating layer and the charge transfer protection film as a second insulating layer;

(h) forming a third conductive layer over the entire surface of the memory region, and then anisotropically etching the third conductive layer into a second control gate on a side surface of each of the first control gates and on the semiconductor layer with the second insulating layer interposed;

(i) forming an impurity layer in the semiconductor layer to form a source region or a drain region;

(j) forming a third insulating layer over the entire surface of the memory region and then removing part of the third insulating layer to expose part of the stopper layer; and

(k) removing the stopper layer, forming a fourth conductive layer over the entire surface of the semiconductor layer, and then patterning the fourth conductive layer to form a word line.

This method of manufacturing a semiconductor device has advantages the same as those of the above embodiment, and enables the control gates to be formed on the insulating layers having different thickness.

Moreover, since the charge transfer protection film is formed in the step (g) after removing the second silicon oxide film of the second insulating layer, the charge transfer protection film can be securely formed on the silicon nitride film.

The method of manufacturing a semiconductor device in accordance with this embodiment could have following features.

(A) In this method of manufacturing a semiconductor device, the charge transfer protection film may be one of a silicon oxide film and a silicon oxide nitride film.

(B) In this method of manufacturing a semiconductor device, the charge

transfer protection film may be formed by a chemical vapor deposition method.

(C) In this method of manufacturing a semiconductor device, the charge transfer protection film may be formed by a thermal oxidation method.

The semiconductor device and the method of manufacturing the same in the embodiment of the present invention are described below in more detail with reference to FIGS. 1 to 17.

1. First embodiment

1.1 Device configuration

FIG. 1 is a plan view showing a layout of a semiconductor device according to the present embodiment. The semiconductor device includes a memory region 1000 including a non-volatile memory device.

In the memory region 1000, MONOS non-volatile memory devices (hereinafter called "memory cells") 100 are arranged in a plurality of rows and columns in the shape of a matrix. A first block B1 and a part of other blocks B0 and B2 adjacent to the first block B1 are illustrated in the memory region 1000. The blocks B0 and B2 have a configuration which is the reverse of the configuration of the block B1.

An element isolation region 300 is formed in a part of a region between the first block B1 and the blocks B0 and B2 adjacent to the first block B1. A plurality of word lines 50 (WL) extending in the X direction (row direction) and a plurality of bit lines 60 (BL) extending in the Y direction (column direction) are provided in each block. One word line 50 is connected with a plurality of word gates 14 arranged in the X direction. The bit lines 60 are formed by impurity layers 16 and 18.

A conductive layer 40 which forms control gates 20 and 30 is formed to enclose each of the impurity layers 16 and 18. Specifically, each of the control gates 20 and 30 extends in the Y direction, and one end of each of a pair of control gates 20 and 30 is connected by the conductive layer extending in the X direction. The other

end of each of the pair of control gates 20 and 30 is connected with one common contact section 200. Therefore, the conductive layer 40 has a function as the control gates of the memory cells and a function as an interconnect which connects each of the control gates arranged in the Y direction.

The single memory cell 100 includes one word gate 14, the control gates 20 and 30, and the impurity layers 16 and 18. The control gates 20 and 30 are formed on each side of the word gate 14. The impurity layers 16 and 18 are formed on the outer side of the control gates 20 and 30. The impurity layers 16 and 18 are shared by the adjacent memory cells 100.

The impurity layer 16 formed in the block B1 and the impurity layer 16 formed in the block B2, adjacent in the Y direction, are electrically connected by a contact impurity layer 400 formed in the semiconductor substrate. The contact impurity layer 400 is formed on the side of the impurity layer 16 opposite to the side of the common contact section 200 of the control gates.

A contact 350 is formed on the contact impurity layer 400. The bit line 60 formed by the impurity layer 16 is electrically connected with an upper interconnect layer through the contact 350.

The impurity layer 18 formed in the block B1 and the impurity layer 18 formed in the block B0, adjacent in the Y direction, are electrically connected by the contact impurity layer 400 on the side on which the common contact section 200 is not disposed. As shown in FIG. 1, the planar layout of a plurality of the common contact sections 200 in one block is in a staggered arrangement in which the common contact sections 200 are alternately formed on opposite ends of the impurity layers 16 and the impurity layers 18. The planar layout of a plurality of the contact impurity layers 400 in one block is in a staggered arrangement in which the contact impurity layers 400 are alternately formed on opposite ends of the impurity layers 16 and the impurity layers 18.

The cross-sectional structure of the semiconductor device is described below with reference to FIGS. 2 and 3. FIG. 2 is a cross-sectional view along the line A-A shown in FIG. 1. FIG. 3 is an enlarged cross-sectional view of a portion B shown in FIG. 2.

In the memory region 1000, the memory cell 100 includes the word gate 14, the impurity layers 16 and 18, and the control gates 20 and 30. The word gate 14 is formed on the semiconductor substrate 10 through a gate insulating layer 12. The impurity layers 16 and 18 are formed in the semiconductor substrate 10. Each of the impurity layers becomes either a source region or a drain region. A silicide layer 92 is formed on the impurity layers 16 and 18.

The control gates 20 and 30 are formed along each side of the word gate 14. The control gate 20 includes a first control gate 20a and a second control gate 20b which are adjacent to each other. The first control gate 20a is formed on the semiconductor substrate 10 through a first insulating layer 22, and is formed on one side surface of the word gate 14 through a side insulating layer 26. The second control gate 20b is formed on the semiconductor substrate through a second insulating layer 24. The control gate 30 includes a first control gate 30a and a second control gate 30b.

The first insulating layer 22 is formed of an ONO film, which is a stack of a bottom silicon oxide layer (first silicon oxide layer) 22a, a silicon nitride layer 22b, and a top silicon oxide layer (second silicon oxide layer) 22c.

The second insulating layer 24 is formed of an ONO film, which is a stack of a bottom silicon oxide layer (first silicon oxide layer) 24a, a silicon nitride layer 24b, and a top silicon oxide layer (charge transfer protection film) 24c. The top silicon oxide layer 24c has a thickness smaller than that of the top silicon oxide layer 22c of the first insulating layer 22.

The first silicon oxide layer 22a forms a potential barrier between a channel region and a charge accumulation region. The silicon nitride layer 22b functions as the

charge accumulation region which traps carriers (electrons, for example). The second silicon oxide layer 22c forms a potential barrier between the control gate and the charge accumulation region.

The side insulating layer 26 is an ONO film. In more detail, the side insulating layer 26 is a stack of a first silicon oxide layer 26a, a silicon nitride layer 26b, and a second silicon oxide layer 26c. The side insulating layer 26 electrically isolates the word gate 14 from each of the control gates 20 and 30. At least the upper end of the first silicon oxide layer 26a of the side insulating layer 26 is located at a position higher than the upper ends of the control gates 20 and 30 with respect to the semiconductor substrate 10 in order to prevent occurrence of short circuits between the word gate 14 and the control gates 20 and 30.

The side insulating layer 26 and the first insulating layer 22 are formed in the same deposition step and have the same layer structure.

The surfaces of the control gates 20 and 30 are covered with a sidewall insulating layer 152.

A buried insulating layer 70 is formed between the adjacent control gates 20 and 30 in the adjacent memory cells 100. The buried insulating layer 70 covers the control gates 20 and 30 so that at least the control gates 20 and 30 are not exposed. The upper surface of the buried insulating layer 70 is located at a position higher than the upper surface of the word gate 14 with respect to the semiconductor substrate 10. The control gates 20 and 30 can be electrically isolated from the word gate 14 and the word line 50 more reliably by forming the buried insulating layer 70 in this manner.

As shown in FIG. 2, the word line 50 is formed on the word gate 14.

In the semiconductor device of the present embodiment, the control gates 20 and 30 respectively include the first control gates 20a and 30a and the second control gates 20b and 30b which are formed on the insulating layers having different thicknesses. Therefore, the potential of the surface of the substrate under the control

gates 20 and 30 varies at two locations, whereby the field intensity has peaks at three locations including the boundary between the word gate 14 and the control gates 20 and 30, the boundary between the first control gates 20a and 30a and the second control gates 20b and 30b, and the edge of the impurity region. This contributes to the following advantages relating to the data write/erase operation of the memory cell 100.

The data write operation is described below. In the case of writing data in the memory cell 100, electrons transferred from the impurity region 16 are provided with energy at the boundary between the word gate 14 and the control gate 30. The electrons are provided with energy at the boundary between the first control gate 30a and the second control gate 30b to become hot electrons. The hot electrons are injected and trapped in the first insulating layer 22 near the region at which the thicknesses of the insulating layers differ.

In the semiconductor device of the present embodiment, electron injection positions are distributed around the boundary between the first control gate 30a and the second control gate 30b. However, since the second insulating layer 24 formed of the NO film is present under the second control gate 30b, the charge escapes through the control gate 30. As a result, electrons trapped on the side of the first control gate 30a remain.

The data erase operation is described below with reference to FIG. 18. FIG. 18 is a band diagram in which the vertical axis indicates electron potential energy and the horizontal axis indicates a real-space coordinate. FIG. 18 shows a state at the edge of the impurity layer 18, specifically, the pn junction.

A high positive voltage is applied to the impurity layer 18 and a negative voltage is applied to the control gate 30. As a result, the electron potential energy is decreased in the impurity layer 18 which is an n-type region (electron potential energy in the n-type region shifts in the direction indicated by an arrow in FIG. 18). Since the thickness of the depletion layer is as small as several nanometers in a high concentration

pn junction, electrons in the p-type valence band can move into the n-type conduction band by a tunneling effect. Specifically, holes are generated near the edge of the impurity layer 18 which is the p-type region accompanying transfer of electrons. This means that a hole accumulation layer is formed near the edge of the impurity layer.

The electric field between the second control gate 30b formed on the second insulating layer 24 and the surface of the substrate and the electric field between the first control gate 30a formed on the first insulating layer 22 and the surface of the substrate are described below. Since the hole accumulation layer is formed in the second insulating layer 24, carrier conductivity of the second insulating layer 24 is high. Therefore, the electric field in the horizontal direction (gate length direction) is relatively small. Since the second insulating layer 24 has a thickness smaller than that of the first insulating layer 22, the electric field in the vertical direction is relatively large. Therefore, holes generated near the edge of the impurity layer 18 cannot enter the second insulating layer 24.

In the first insulating layer 22, the electric field in the horizontal direction is relatively large and the electric field in the vertical direction is relatively small. Therefore, holes generated near the edge of the impurity layer 18 are provided with a large amount of energy at the boundary between the second insulating layer 24 and the first insulating layer 22 and enter the charge accumulation film. Specifically, holes are injected at a location near the region in which the thicknesses of the charge accumulation films differ, whereby data is erased at this location.

The location at which electrons are injected during writing can be allowed to coincide with the location at which holes are injected during erasing in this manner. As a result, a non-volatile memory device which does not deteriorate even if the write/erase cycles are repeated can be realized.

Moreover, according to the semiconductor device of the present embodiment, the charge transfer protection film is formed of the silicon oxide film 24c in the

uppermost layer of the second insulating layer 24. Therefore, the silicon nitride films 22b and 24b in which a charge is accumulated can be prevented from coming in contact with the second control gate 30b. This prevents the charge in the silicon nitride films 22b and 24b from diffusing into the second control gate 30b, whereby charge retention characteristics can be improved.

1.2 Method of manufacturing semiconductor device

A method of manufacturing the semiconductor device according to the present embodiment is described below with reference to FIGS. 4 to 14. Each cross-sectional view corresponds to the section along the line A-A shown in FIG. 1. In FIGS. 4 to 14, sections the same as those shown in FIGS. 1 to 3 are indicated by the same symbols. The description which has already been given is omitted.

(1) The element isolation region 300 (see FIG. 1) is formed on the surface of the semiconductor substrate 10 by using a trench isolation method. p-type impurities are implanted as channel doping. The n-type contact impurity layer 400 (see FIG. 1) is formed in the semiconductor substrate 10 by ion implantation.

As shown in FIG. 4, an insulating layer 120 which becomes the gate insulating layer is formed on the surface of the semiconductor substrate 10. A gate layer (first conductive layer) 140 which becomes the word gate 14 is deposited on the first insulating layer 120. The gate layer 140 is formed of doped polysilicon. A stopper layer S100 used in a CMP step described later is formed on the gate layer 140. The stopper layer S100 is formed of a silicon nitride layer.

(2) A resist layer (not shown) is formed. The stopper layer S100 is patterned by using the resist layer as a mask. The gate layer 140 is etched by using the patterned stopper layer as a mask. As shown in FIG. 5, the gate layer 140 is patterned to become a gate layer (word gate) 140a.

FIG. 6 is a plan view showing a state after patterning. The patterning allows

openings 160 and 180 to be formed in the laminate of the gate layer 140a and the stopper layer S100 in the memory region 1000. The openings 160 and 180 approximately correspond to the regions in which the impurity layers 16 and 18 are formed by ion implantation described later. The side insulating layer and the control gate are formed along the side surfaces of the openings 160 and 180 in a step described later.

(3) The surface of the semiconductor substrate is washed by using diluted fluoric acid. This allows the insulating layer 120 to be removed in the exposed region, whereby the gate insulating layer 12 remains. As shown in FIG. 7, a first silicon oxide layer 220a is deposited by using a thermal oxidation method. The first silicon oxide layer 220a is formed on the exposed surface of the semiconductor substrate 10 and the gate layer 140a. The first silicon oxide layer 220a may be formed by using a CVD method.

The first silicon oxide layer 220a is subjected to an annealing treatment. The annealing treatment is performed in an atmosphere containing NH_3 gas. This pretreatment enables a silicon nitride layer 220b to be uniformly and easily deposited on the first silicon oxide layer 220a. The silicon nitride layer 220b is deposited by using a CVD method.

A second silicon oxide layer 220c is deposited by using a CVD method, specifically, a high temperature oxidation (HTO) method. The second silicon oxide layer 220c may be deposited by using an IN-situ steam generation (ISSG) treatment. A dense film is deposited by using the ISSG treatment. In the case of depositing the second silicon oxide layer 220c by using the ISSG treatment, an annealing treatment for making the ONO film dense as described layer may be omitted.

In the above step, since the silicon nitride layer 220b and the second silicon oxide layer 220c are deposited in the same furnace, interfacial contamination due to removal from the furnace can be prevented. This enables a uniform ONO film to be

formed, whereby the memory cell 100 having stable electrical characteristics can be obtained.

In the present embodiment, the ONO film 220 becomes the first insulating layer 22, the second insulating layer 24, and the side insulating layer 26 (see FIG. 2) by patterning described later.

(4) As shown in FIG. 8, a doped polysilicon layer (second conductive layer) 230 is formed on the second silicon oxide layer 220c. The doped polysilicon layer 230 is etched in a subsequent step to become the conductive layer 40 (see FIG. 1) which forms the control gates 20 and 30.

(5) As shown in FIG. 9, the entire surface of the doped polysilicon layer 230 is anisotropically etched. This allows a sidewall-shaped conductive layer 232 to be formed along the side surfaces of the openings 160 and 180 (see FIG. 5) in the memory region 1000. The sidewall-shaped conductive layer 232 is etched in a step described later and becomes the first control gates 20a and 30a.

(6) The surface of the second silicon oxide layer 220c of the ONO film 220 is removed by using the sidewall-shaped conductive layer 232 as a mask. In more detail, the second silicon oxide layer 220c may be removed by wet etching using diluted fluoric acid, or dry etching. This allows the first insulating layer 22 consisting of the ONO film to remain under the first control gates 20a and 30a. As shown in FIG. 10, the second silicon oxide layer 220c is etched while allowing a part of the second silicon oxide layer 220c to remain so that the silicon nitride film 220b is not exposed. The etched ONO film 220 becomes the second insulating layer consisting of the first silicon oxide layer 24a, the silicon nitride layer 24b, and the second silicon oxide layer (charge transfer protection film) 24c in a step described later.

(7) A doped polysilicon layer (not shown) is formed over the entire surface. The entire surface of the doped polysilicon layer is anisotropically dry-etched. This allows the first control gates 20a and 30a to be formed by decreasing the height of the

sidewall-shaped conductive layer 232, and the second control gates 20b and 30b to be formed on the stack of the second insulating layer 24 consisting of the first silicon oxide layer 24a, the silicon nitride layer 24b, and the second silicon oxide layer 24c, as shown in FIG. 11.

Since the first control gates 20a and 30a and the second control gates 20b and 30b are formed in the same step, the heights of the first control gates 20a and 30a and the second control gates 20b and 30b can be easily made uniform. The surfaces of the control gates 20 and 30 are gently sloped by isotropic etching. This allows the exposed second silicon oxide layer 24c to be removed.

(8) An insulating layer (not shown) such as silicon oxide or silicon nitride oxide is formed over the entire surface of the memory region 1000. As shown in FIG. 12, the sidewall insulating layer 152 is formed to cover the control gates 20 and 30 by anisotropically etching the insulating layer. The insulating layer deposited in a region in which the silicide layer is formed in a step described later is removed by this etching, whereby the semiconductor substrate is exposed.

As shown in FIG. 12, the impurity layers 16 and 18 are formed in the semiconductor substrate 10 by ion implantation with n-type impurities.

A metal for forming a silicide is deposited over the entire surface. As examples of the metal for forming a silicide, titanium, cobalt, and the like can be given. The silicide layer 92 is formed on the exposed surface of the semiconductor substrate by subjecting the metal formed on the semiconductor substrate to a silicidation reaction. A third insulating layer 270 such as silicon oxide or silicon nitride oxide is formed over the entire surface of the memory region 1000. The third insulating layer 270 is formed to cover the stopper layer S100.

(9) As shown in FIG. 13, the third insulating layer 270 is planarized by grinding the third insulating layer 270 by using a CMP method until the stopper layer S100 is exposed. This grinding allows the buried insulating layer 70 to remain

between the control gates 20 and 30 which face each other.

(10) The stopper layer S100 is removed by using thermal phosphoric acid. As a result, at least the upper surface of the gate layer 140a is exposed, whereby an opening 170 is formed in the buried insulating layer 270. Specifically, the opening 170 is a region which is formed by removing the stopper layer S100 and is located on the gate layer 140a.

(11) A doped polysilicon layer (not shown) is deposited over the entire surface. A patterned resist layer (not shown) is formed on the doped polysilicon layer. The doped polysilicon layer is patterned by using the resist layer as a mask, whereby the word line 50 is formed. The gate layer 140a (see FIG. 14) is etched by using the resist layer as a mask. This allows the gate layer 140a to be removed in the region in which the word line 50 is not formed. As a result, the word gates 14 (see FIG. 1) arranged in an array are formed. The region in which the gate layer 140a is removed corresponds to the region in which the p-type impurity layer (element isolation impurity layer) 15 is formed later.

Since the first and second control gates 20 and 30 are covered with the buried insulating layer 70 in this etching step, the control gates 20 and 30 remain without being etched.

The entire surface of the semiconductor substrate 10 is doped with p-type impurities. This allows the p-type impurity layer (element isolation impurity layer) 15 (see FIG. 1) to be formed in the region between the word gates 14 adjacent in the Y direction. The non-volatile memory devices 100 can be isolated more reliably by the p-type impurity layer 15.

The semiconductor device shown in FIGS. 1 to 3 is manufactured by these steps.

Advantages of this manufacturing method are as follows.

The control gates 20 and 30 are formed in two steps. In more detail, the first

control gates 20a and 30a are formed. The second silicon oxide layer 220c of the ONO film 220 is removed. The second control gates 20b and 30b are then formed. Therefore, the control gates 20 and 30 can be formed on the insulating layers having different thicknesses. As a result, a semiconductor device in which field intensity between the control gates 20 and 30 and the surface of the substrate is nonuniform can be manufactured.

In the step (6), the surface of the ONO film 220 is removed by using the sidewall-shaped conductive layer 232 as a mask. Specifically, the ONO film 220 is etched so that a part of the second silicon oxide film 220c remains on the silicon nitride film 220b. Therefore, the silicon oxide film which functions as the charge transfer protection film can be allowed to remain on the silicon nitride film 220b. As a result, the silicon nitride layer 22b of the first insulating layer 22 can be prevented from coming in contact with the second control gates 20b and 30b, whereby a semiconductor device having improved charge retention characteristics can be manufactured.

2. Second embodiment

A second embodiment of the present invention is described below. The following description merely illustrates features differing from those of the first embodiment.

2.1 Device configuration

FIGS. 15 and 16 are a cross-sectional views schematically showing a semiconductor device according to the second embodiment. FIGS. 15 and 16 are a cross-sectional views showing a portion corresponding to FIG. 3 in the first embodiment. In the semiconductor device according to the second embodiment, a charge transfer protection film 42 is formed in the uppermost layer of the second insulating layer 24.

In the semiconductor device shown in FIG. 15, the second control gate 30b is formed on the first control gate 30a through the charge transfer protection film 42, and is formed on the semiconductor substrate 10 through the second insulating layer 24 including the charge transfer protection film 42 in the uppermost layer. There are no specific limitations to the charge transfer protection film 42 insofar as the charge transfer protection film 42 has a function of preventing a charge injected into the silicon nitride films 22b and 24b from being discharged to the second control gate 30b. For example, a silicon oxide film may be used as the charge transfer protection film 42. It is preferable that the thickness of the charge transfer protection film 42 be smaller than the thickness of the first silicon oxide film 22c of the first insulating layer 22. In the semiconductor device shown in FIG. 15, the charge transfer protection film 42 is formed by using a CVD method. In this case, the charge transfer protection film 42 is formed to cover the silicon nitride film 24b of the second insulating layer 24, the first control gate 30a, and the side insulating layer 26.

In the semiconductor device shown in FIG. 16, the second control gate 30b is formed on the first control gate 30a through the charge transfer protection film 42, and is formed on the semiconductor substrate 10 through the second insulating layer 24 including the charge transfer protection film 42 in the uppermost layer. There are no specific limitations to the charge transfer protection film 42 insofar as the charge transfer protection film 42 has the above-described function. For example, a silicon oxide nitride film may be used as the charge transfer protection film 42. It is preferable that the thickness of the charge transfer protection film 42 be smaller than the thickness of the first silicon oxide film 22c of the first insulating layer 22. In the semiconductor device shown in FIG. 16, the charge transfer protection film 42 is formed by using a thermal oxidation method. In this case, the charge transfer protection film 42 is formed on the silicon nitride film 24b of the second insulating layer 24, the first control gate 30a.

In the semiconductor device according to the second embodiment, the charge transfer protection film 42 is formed of an oxide film or a silicon oxide nitride film in the uppermost layer of the second insulating layer 24. Therefore, the silicon nitride films 22b and 24b can be prevented from coming in contact with the second control gate 20b. This prevents electrons accumulated in the silicon nitride films 22b and 24b from being discharged to the second control gate 20b, whereby a semiconductor device having improved charge retention characteristics can be provided.

2.2 Method of manufacturing semiconductor device

A method of manufacturing the semiconductor device according to the second embodiment is described below.

The steps (1) to (5) are performed in the same manner as in the first embodiment.

(6) The second silicon oxide layer 220c of the ONO film 220 is removed by using the sidewall-shaped conductive layer 232 as a mask. This allows the first insulating layer 22 formed of the ONO film to remain under the first control gates 20a and 30a. In more detail, the second silicon oxide layer 220c may be removed by wet etching using diluted fluoric acid, or dry etching. The charge transfer protection film 42 (not shown) is formed over the entire surface. A silicon oxide film or a silicon oxide nitride film may be formed as the charge transfer protection film 42. The charge transfer protection film 42 may be formed by using a CVD method or a thermal oxidation method.

(7) A doped polysilicon layer (not shown) is formed over the entire surface. The entire surface of the doped polysilicon layer is anisotropically dry-etched. This allows the first control gates 20a and 30a to be formed by decreasing the height of the sidewall-shaped conductive layer 232, and the second control gates 20b and 30b to be formed on the stack of the second insulating layer 24 consisting of the first silicon oxide

layer 24a, the silicon nitride layer 24b, and the charge transfer protection film 42, as shown in FIG. 10.

The steps (8) to (11) are then performed in the same manner as in the first embodiment to obtain the semiconductor device shown in FIG. 15.

According to the manufacturing method of the present embodiment, the charge transfer protection film 42 is formed in the step (6) after removing the second silicon oxide film 220c. Therefore, the second control gate 30b can be formed on the second insulating layer 24 consisting of the silicon oxide layer 24a, the silicon nitride layer 24b, and the charge transfer protection film 42. As a result, a semiconductor device in which the charge retention characteristics of the silicon nitride films 22b and 24b are improved can be manufactured.

The present invention is not limited to the above-described embodiments. Various modifications and variations are possible within the scope of the present invention. For example, a bulk semiconductor substrate is used as the semiconductor layer in the above embodiments. However, a semiconductor layer of an SOI substrate may be used.